

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2003, 23:55:25 ; Search time 6291 Seconds
(without alignments)
9918.362 Million Cell updates/sec

Title: US-09-698-781-2

Perfect score: 2144

Sequence: 1 tgatqaacaatacttcat.....aaaaaaaaaaaaaaaa 2144

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : GenEmpl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: go_v1:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_lov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_v1:*

29: em_vt:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rdn:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htg_hum:*

40: em_htg_mus:*

41: em_htg_other:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2144	100.0	2144	6	AX127587	AX127587 Sequence
2	2134.8	99.6	2138	9	HSSP032	X94223 H. sapiens m
3	2113.2	98.6	2128	6	AX335634	AX335634 Sequence
4	2113.2	98.6	2128	9	HSCR1SP3G	X9240 H. sapiens m
5	1485.2	69.3	151752	2	AC010779	AL01950 Human DNA
6	1480.4	69.0	118524	9	HSJD42L6	M2532 Human testi
7	515.2	24.0	1335	9	HUMTP1A	BC022011 Homo sapi
8	513.6	24.0	1380	9	BC022011	X92239 H. sapiens m
9	513.6	24.0	1406	9	HSCR1SP2I	AJ001400 Equus cab
10	505.8	23.6	1295	4	ECCR1SP3	U55712 Cavia porce
11	476.2	22.2	1388	10	CPU35712	M25533 Mouse testi
12	370.8	17.3	1418	10	MUSTX1A	AB009662 Rattus no
13	367.8	17.2	1280	10	AB009662	AF078552 Rattus no
14	362.2	16.9	1498	10	RATBG	M31173 Rat epididy
15	362.2	16.9	1498	10	MUSAG1A	M2849 Mouse acid
16	346.4	16.2	1380	10	MUSCR1SP	L05559 Mouse cyste
17	346.4	16.2	1403	10	BC011150	BC011150 Mus mus
18	341.6	15.9	1445	10	RNESCR	X0643 Ratt mRna fo
19	325.6	15.2	928	10	ECA6632	AJ006532 Equus cab
20	325.4	15.2	495	4	BTA277708	AJ277708 Bos tauru
21	294.4	13.7	453	4	MUSAG2A	M2850 Mouse acid
22	287.6	13.4	1394	10	BC022573	BC022573 Mus mus
23	287.6	13.4	1418	10	MUSC1SPB	L05560 Mouse cyste
24	287	13.4	1406	10	AX127586	AX127586 Sequence
25	254.4	11.9	2556	6	U13619	U13619 Heloderma h
26	226.4	10.6	1090	10	U13619	AJ277709 Bos tauru
27	210.8	9.8	334	4	BTA277709	AF384218 Agkistrod
28	205	9.6	1336	5	AF384218	AF384218 Agkistrod
29	202.4	9.4	1321	5	ECA315379	AF384219 Trimeresurus
30	198.6	9.3	791	5	AY039355	AY039355 Rhabdophi
31	198.6	9.3	1318	5	PMU59447	U59447 Trimeresurus
32	183	8.5	51371	9	AL353193	AL353193 Human DNA
33	182.8	8.5	169189	2	AC124008	AC124008 Equus cab
34	182.4	8.5	1309	5	AF190561	AF190561 Equus cab
35	182	8.5	1475	4	ECA315379	AJ315379 Equus cab
36	180.8	8.4	1305	2	AF384220	AF384220 Equus cab
37	179.8	8.4	15869	2	AC023420	AC023420 Homo sapi
38	177.4	8.3	1316	5	AF159541	AF159541 Lapemis h
39	177	8.3	1281	9	D38451	D38451 Homo sapi
40	177	8.3	1282	9	S80310	S80310 acidic epid
41	177	8.3	1886	9	HSCR1SP1G	X95237 H. sapiens m
42	175.6	8.2	1174	4	ECA6631	AJ006631 Equus cab
43	173.4	8.1	149907	4	AC091436	AC091436 Felis cat
44	170.2	7.9	166287	9	AC026643	AC026643 Homo sapi
45	169.2	7.9	1343	9	AF123894	AF123894 Macaca mu

ALIGNMENTS

RESULT	1	AX127587	AX127587	Sequence 2	from Patent WO0131343.	DNA	linear	PAT 15-MAY-2001
DEFINITION								
ACCESSION								
VERSION								
KEYWORDS								
SOURCE								
ORGANISM								
HUMAN								
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.								
REFERENCE	1	(bases 1 to 2144)						
AUTHORS	Hubert,R.S., Raitano,A.B., Afar,D.E., Mitchell,S.C., Farts,M. and Jakobovits,A.							
TITLE	Diagnosis and therapy of cancer using sgp28-related molecules							

Pred. No. is the number of results predicted by chance to have a

JOURNAL	Patent: WO 0131343-A 2 03-MAY-2001;
FEATURES	Urogenesys, Inc. (US) Location/Qualifiers
source	1. 2144 /organism="Homo sapiens" /db_xref="taxon: 9606"
BASE COUNT	735 a 403 c 382 g 624 t
ORIGIN	
Query Match	100 %; Score 2144; DB 6; Length 2144;
Best Local Similarity	100 %; Pred. No. 0;
Matches	2144; Conservative 0; Mismatches 0;
Db	Indels 0; Gaps 0;
Qy	1 TGATGAAACAATACATCCTGCTCTGGAAACACTGCATGACATATTCCAGTG 60
Db	1 TGATGAAACAATACATCCTGCTCTGGAAACACTGCATGACATATTCCAGTG 60
Qy	61 TGTGTTCCCTGGTGGCTGGCTGGCTGCCTCCATCTTTCAGCAATGAGATAGGATCCG 120
Db	61 TGTGTTCCCTGGTGGCTGGCTGCCTCCATCTTTCAGCAATGAGATAGGATCCG 120
Qy	121 CTTTACTGCCTTGTAACCAACCCAAACACAGTCACAAGGGAGTTGGAAATAAGCACA 180
Db	121 CTTTACTGCCTTGTAACCAACCCAAACACAGTCACAAGGGAGTTGGAAATAAGCACA 180
Qy	181 ATGAACTGAGGAGCAGTATCCTCCCTGCCAGAACATGCTGAGANGGATGAC 240
Db	181 ATGAACTGAGGAGCAGTATCCTCCCTGCCAGAACATGCTGAGANGGATGAC 240
Qy	241 AAGGGCTGCAGCAATGCCAAGTGGCAACAGTCGCAATTACAGACACTAAC 300
Db	241 AAGAGCTSCAGCAATGCCAAGTGGCAACAGTCGCAATTACAGACACTAAC 300
Qy	301 CAAGGATCCGAAATGACAAGCTCTAAATGTTGAGCAATCTACATGTCAGTGCCCCA 360
Db	301 CAAGGATCCGAAATGACAAGCTCTAAATGTTGAGCAATCTACATGTCAGTGCCCCA 360
Qy	361 GCTCAGGTCACAGCAATCCAAACGCTGGTTGATGAGTCATGTTG 420
Db	361 GCTCAGGTCACAGCAATCCAAACGCTGGTTGATGAGTCATGTTG 420
Qy	421 TAGGCCAAGAGCTCCAAAGCAGTGGTGGACATATACACGCTGTTGGACTCT 480
Db	421 TAGGCCAAGAGCTCCAAAGCAGTGGTGGACATATACACGCTGTTGGACTCT 480
Qy	481 CATACTCGTGTGATGTTGAAATGCTACTGTCCTCAAAAGTCTAAACTACT 540
Db	481 CATACTCGTGTGATGTTGAAATGCTACTGTCCTCAAAAGTCTAAACTACT 540
Qy	541 ATGTTGCCATATGTCCTGTTAATGGCTAATGACTATATGTCCTTATGAC 600
Db	541 ATGTTGCCATATGTCCTGTTAATGGCTAATGACTATATGTCCTTATGAC 600
Qy	601 AAGGGCACCTTGGCCAGATRACTGTGACCATGACTATGCCAATGGT 660
Db	601 AAGGGCACCTTGGCCAGATRACTGTGACCATGACTATGCCAATGGT 660
Qy	661 GCAAGTACCAAGATCTATAGTACTGAAATTTGAGCTCACATTAACCTGAAC 720
Db	661 GCAAGTACCAAGATCTATAGTACTGAAATTTGAGCTCACATTAACCTGAAC 720
Qy	721 ATCACTGGTCAGGAGCAGTGTGCAAGGATCTGAAACTGTCATGAAAC 780
Db	721 ATCACTGGTCAGGAGCAGTGTGCAAGGATCTGAAACTGTCATGAAAC 780
Qy	781 TACGGATTGACCCAGGAGCTATGAGGAGGAGTCAGTATCTACTAGATTG 840
Db	781 TACGGATTGACCCAGGAGCTATGAGGAGGAGTCAGTATCTACTAGATTG 840
Qy	841 GCACTACTAGTATGACTATAGTACTGAGAAATCTGAGCATGTTGACACATT 900
Db	841 GCACTACTAGTATGACTATAGTACTGAGAAATCTGAGCATGTTGACACATT 900
Qy	901 TGATTCAAATGTTCTCTGGATCTGCCTTATTACAAAATTTTCATACA 960
Db	901 TGATTCAAATGTTCTCTGGATCTGCCTTATTACAAAATTTTCATACA 960
Qy	961 AATGGTTAAAGAACAACAACTCTAAACACAACACTTGGATTATATAACTTG 1020
Db	961 AATGGTTAAAGAACAACAACTCTAAACACAACACTTGGATTATATAACTTG 1020
Qy	1021 TGATTAATTACTGAAATTATAGGGAAATTIGAACACTGTTCTCATGA 1080
Db	1021 TGATTAATTACTGAAATTATAGGGAAATTIGAACACTGTTCTCATGA 1080
Qy	1081 CTAAGTCTACTAACCCCTGGATGAGGAAATTGTCCTAGACAAATGTA 1140
Db	1081 CTAAGTCTACTAACCCCTGGATGAGGAAATTGTCCTAGACAAATGTA 1140
Qy	1141 AAAGACAAATAATTACATGACCCCTGGCTGACTGTGCTTCTAGTCAC 1200
Db	1141 AAAGACAAATAATTACATGACCCCTGGCTGACTGTGCTTCTAGTCAC 1200
Qy	1201 CTAAGCTAACGCACTCTCAAGACGTTCCATATGCTGCTTAATTTCACCT 1260
Db	1201 CTAAGCTAACGCACTCTCAAGACGTTCCATATGCTGCTTAATTTCACCT 1260
Qy	1261 TCACCTTCTCCCATCTCTGGCATCTCACATGTCAGTGAACCTGTCCTC 1320
Db	1261 TCACCTTCTCCCATCTCTGGCATCTCACATGTCAGTGAACCTGTCCTC 1320
Qy	1321 TAAACATCTGACTTTATTCGCAAACTACATACATCCTTGAATTTTATCTG 1380
Db	1321 TAAACATCTGACTTTATTCGCAAACTACATACATCCTTGAATTTTATCTG 1380
Qy	1381 CATAATTTCAGTAGAAATGATGATGAACTCTCATTTAACCTCTTGTG 1440
Db	1381 CATAATTTCAGTAGAAATGATGATGAACTCTCATTTAACCTCTTGTG 1440
Qy	1441 AAACTCTCTAAAGAATACAGATAATAGGTAATACCTCCACTCAGGG 1500
Db	1441 AAACTCTCTAAAGAATACAGATAATAGGTAATACCTCCACTCAGGG 1500
Qy	1501 AGAACTCAGTCCTCCCTGTGACTCTACTAAATGACTCTCCAGTCAGGT 1560
Db	1501 AGAACTCAGTCCTCCCTGTGACTCTACTAAATGACTCTCCAGTCAGGT 1560
Qy	1561 GGAGTAGGAGGAAAGATAGTAACTTACAGGGAGAAATGACAAATGAGCT 1620
Db	1561 GGAGTAGGAGGAAAGATAGTAACTTACAGGGAGAAATGACAAATGAGCT 1620
Qy	1621 CACCAAGTGTCAAAATTACGTCAACAGTGTGACATTCAGTTCTGATATA 1680
Db	1621 CACCAAGTGTCAAAATTACGTCAACAGTGTGACATTCAGTTCTGATATA 1680
Qy	1681 CTTCCTAAATTCATACATCCAACTCTAAATTAGGTTAGCTAACCTCAGT 1740
Db	1681 CTTCCTAAATTCATACATCCAACTCTAAATTAGGTTAGCTAACCTCAGT 1740
Qy	1741 TGAGGACATCTCAAAATCTCTGGTATTAGGTTAGCTAACACTGAA 1800
Db	1741 TGAGGACATCTCAAAATCTCTGGTATTAGGTTAGCTAACACTGAA 1800
Qy	1801 ATCACTGAAATAGGGAACTCTGAGGAGTCAGTATCTGAGCAACACTGAA 1860
Db	1801 ATCACTGAAATAGGGAACTCTGAGGAGTCAGTATCTGAGCAACACTGAA 1860
Qy	1861 TGTGAGCCAAATGCACTGCTCTTGGTCAATCTGAGAAACATCACAGCTAGT 1920
Db	1861 TGTGAGCCAAATGCACTGCTCTTGGTCAATCTGAGAAACATCACAGCTAGT 1920
Qy	1921 GAAAACCTGATGACTGAAATCTGAGAAATCTGAGCAACACTGAGCTAGT 1980
Db	1921 GAAAACCTGATGACTGAAATCTGAGAAATCTGAGCAACACTGAGCTAGT 1980
Qy	1981 TAATCTGACAATATAGGGAAATGTAAGATGATAACGTTAGAGAAACTGG 2040

TITLE		Cancer gene determination and therapeutic screening using signature gene sets	
JOURNAL		Patent: WO 0194629-A 6143 13-DEC-2001; Avalon Pharmaceuticals (US), Location/Qualifiers	
FEATURES	source	1. .2128 /organism="Homo sapiens" /db_xref="Taxon:9606"	
BASE COUNT	734 a	397 c	380 g
ORIGIN			
Query	Match	98.6%	Score 2113.2;
Matches	Local Similarity	99.9%	DB 6;
2115;	Conservative	0;	Length 2128;
0;	Mismatches	3;	Indels 0;
0;	Gaps 0;	0;	
Oy	27	CTGGAAACACTGCCAATGACATATTCCAGTGTGCTTACGCTGCTCT 86	
Db	1	CTGGAAACACTGCCAATGACATATTCCAGTGTGCTTACGCTGCTCT 60	
Oy	87	CCATCTTCCAGCAATGAGTAAGATCCGGCTTACGCTGCTCT 146	
Db	61	CCATCTTCCAGCAATGAGTAAGATCCAGTGTGCTTACGCTGCTCT 120	
Oy	147	ACACAGTGCAGGAGATGGAATAGCAGATGAGCTGAGCTACTCC 206	
Db	121	ACACAGTGCAGGAGATGGAATAGCAGATGAGCTGAGGAGAGCAGTACTCC 180	
Oy	1682	TTCTAAATTCATACTCCAACTTAATGAGCTAACATCCAGGAACTCAAGT 1741	
Db	1681	TTCTCTAAATTCATACTCCAACTTAATGAGCTAACATCCAGGAACTCAAGT 1740	
Oy	1742	GAGGAGATCTACAAATATCCTGGGTTTTAGAGTATCTCAAACCTAA 1801	
Db	1741	GAAGGAGATCTACAAATATCCTGGGTTTTAGAGTATCTCAAACCTAA 1800	
Oy	1802	TCATGGAAATAGGGATCTGAGAACAACTACAGACACATGAGACTAAGGAC 1861	
Db	1801	TCATGGAAATAGGGATCTGAGAACAACTACAGACACATGAGACTAAGGAC 1860	
Oy	1862	GTGAGCCAAATGCAATGCTCTGGATCAGATCCTGGAGAACAGATCGTAATG 1921	
Db	1861	GTGAGCCAAATGCAATGCTCTGGATCAGATCCTGGAGAACAGATCGTAATG 1920	
Oy	1922	AAAACATGAGCTGGAATGAGCTGGAGATTTAACATGAGCTGTTCT 1981	
Db	1921	AAAACATGAGCTGGAATGAGCTGGAGATTTAACATGAGCTGTTCT 1980	
Oy	1982	AACTTGACAAATAGGGTAATGAGTAACTAGCTGAGAACACTGAATG 2041	
Db	1981	AACTTGACAAATAGGGTAATGAGTAACTAGCTGAGAACACTGAATG 2040	
Oy	2042	TGAGGGCTATCTAGGAATCTCTGACTATCTTACCAAAATTCTGGTAGTCTAGAAG 2101	
Db	2041	TGAGGGCTATCTAGGAATCTCTGACTATCTTACCAAAATTCTGGTAGTCTAGAAG 2100	
Oy	2102	CAATGCAAATAAAAGTACTCTGAAAAA 2139	
Db	2101	CAATGCAAATAAAAGTCTGAAAAA 2138	
RESULT 3			
LOCUS	AX35634	AX35634	2128 bp DNA
DEFINITION	Sequence 6143 from Patent WO0194629.	linear	PAT 09-JAN-2002
ACCESSION	AX35634		
VERSION	AX35634.1	GI:18126353	
KEYWORDS	human		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. ¹		
REFERENCE	Young, P.E., Augustus M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppet, D.R. and Weaver, Z.		

Db	841	ACGTGAGAATTTGGAGTGTGATACACATTGTCATAAAGTTCTCTGGT	900
Qy	927	CTCTTTTACAAATTTTCTACAAATGCTAAAGAACAAATCT	986
Db	901	CTGCTTATTACAAATAATTTCATACAAATGCTAAAGAACAAATCT	960
Qy	987	ACACAACTTGGATTATATAAACCTTGTGATTAAATTCTGATTAA	1046
Db	961	ANCAACACTTGGATTATATAAACCTTGTGATTAAATTCTGATTAA	1020
Qy	1047	GGGTGAATTTGAAAGTGTGATTCTCATGACTAACTTCAAACTGATTG	1106
Db	1021	GGGTGAATTTGAAAGTGTGATTCTCATGACTAACTTCAAACTGATTG	1080
Qy	1107	AAGTGAATTTGTCAGACAAAGTACAAAGACAAATATTCAGA	1166
Db	1081	ACGTGAATTTGTCAGACAAAGTACAAAGACAAATATTCAGA	1140
Qy	1167	ACCTTGGTGTAGTGCCTCTAGCTCCACTCTAACGCTAAGCATG	1226
Db	1141	ACCTTGGTGTAGTGCCTCTAGCTCCACTCTAACGCTAAGCATG	1200
Qy	1227	TTGCCATATGCTCTTATCTCTACTCATCACCTCACCTAACCTGGT	1286
Db	1201	TTGCCATATGCTCTTATCTCTACTCATCACCTCACCTAACCTGGT	1260
Qy	1287	GSCATCCACAACTGAGCTGCTCTAAACAACTGCTGACTTATTTG	1346
Db	1261	GGCATCCACAACTGAGCTGCTCTAAACAACTGCTGACTTATTTG	1320
Qy	1347	CAAATCATACATCCTGAAATTTCATGCTACATGAAATTGACATG	1406
Db	1321	CAAATCATACATCCTGAAATTTCATGCTACATGAAATTGACATG	1380
Qy	1407	AAACCTGATTTAACCTCTCTCTGACAAACCTCCCTAAAGAATCAG	1466
Db	1381	AAACCTGATTTAACCTCTCTGACAAACCTCCCTAAAGAATCAG	1440
Qy	1467	ATATATAGTAAATACCTCCACTCAAGAGGAGAACTCAGTCCTCTGTG	1526
Db	1441	ATATATAGTAAATACCTCCACTCAAGAGGAGAACTCAGTCCTCTGTG	1500
Qy	1527	CTCACTAAATGACTCACTCCAAAGAGGAGATGGATGGAAAGGAGAACATG	1586
Db	1501	CTTCACTAAATGACTCACTCCAAAGAGGAGATGGATGGAAAGGAGAACATG	1560
Qy	1587	CTTACAGGGAGAAATGACAATGAGGTCTCACCAAGTGTCAAATTAGTC	1646
Db	1561	CTTACAGGGAGAAATGACAATGAGGTCTCACCAAGTGTCAAATTAGTC	1620
Qy	1647	CAGTGTAGTCATCAGATTGCTACATAATCTTAAATCTAAATCCATC	1706
Db	1621	CAGTGTAGTCATCAGATTGCTACATAATCTTAAATCTAAATCCATC	1680
Qy	1707	TATTTGGCTAAACCTCCACAAACTCAAGTGTAAAGACATCTCAAATTC	1766
Db	1681	TAATATGGCTAAACCTCCACAAACTCAAGTGTAAAGACATCTCAAATTC	1740
Qy	1767	GGGTATTAGTAAACATCCACACTCAAGTGTAAAGACATCTCAAATTC	1826
Db	1741	GGGTATTAGTAAACATCCACACTCAAGTGTAAAGACATCTCAAATTC	1800
Qy	1827	AAACAACTCAGACCACTGAGACTAAGGAGACATGGCCAAATGCAATG	1886
Db	1801	AAACAACTCAGACCACTGAGACTAAGGAGACATGGCCAAATGCAATG	1860
Qy	1887	GGATCAGATCTGGACAGAAAGATGCTAACTGAAACTGATGAGCTGT	1946
Db	1861	GGATCAGATCTGGACAGAAAGATGCTAACTGAAACTGATGAGCTGT	1920
Qy	1947	ATCGAGATTCTGGACAGAAAGATGCTAACTGAAACTGATGAGCTGT	2006
Db	1921	ATCGAGATTCTGGACAGAAAGATGCTAACTGAAACTGATGAGCTGT	1980
Qy	2007	TGAGATGATAAGCTGAGAACACTGGTGGGGCTCTCTGAGATCTCTGT	2056
Db	1981	TGAGATGATAAGCTGAGAACACTGGTGGGGCTCTCTGAGATCTCTGT	2040
Qy	2067	ACTCTTACAAATTGGTAGCTAGAACGACTTAAAGTCTCTGAA	2126
Db	2041	ACTCTTACAAATTGGTAGCTAGAACGACTTAAAGTCTCTGAA	2100
Qy	2127	AAAAAAAALAAAAA 2144	2118
Db	2101	AAAAAAAALAAAAA 2118	
RESULT 4			
REFERENCE	HSCRISPG	HSRISPG	
AUTHORS	H.sapiens	mRNA for cysteine-rich secretory protein-3.	
DEFINITION			
ACCESSION	X95240		
VERSION	X95240.1	GR:1162818	
KEYWORDS	CRISP-3 gene; cysteine-rich secretory protein-3.		
SOURCE	Homo sapiens.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 2128)		
AUTHORS	Kratschmar,J., Haendler,B., Eberspacher,U., Roosterman,D., Donner,P. and Schleuning,W.D.		
TITLE	The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3		
JOURNAL	Eur.J. Biochem. 236 (3), 827-836 (1996)		
PUBMED	96270732		
REFERENCE	2 (bases 1 to 2128)		
AUTHORS	Haendler,B.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-JAN-1996) B. Haendler, Schering AG, ICBM, S109/517, 13342 Berlin, FRG		
FEATURES	source	Location/Qualifiers	
gene		1..2128	
CDS		/organism="Homo sapiens"	
		/ab_xref="taxon:9606"	
		/clone_lib="testis" CDNA library"	
		16..753	
		/ab_xref="CRISP-3"	
		16..753	
		/gene="CRISP-3"	
		/codon_start=1	
		/product="cysteine-rich, secretory protein-3"	
		/protein_id="CA64527.1"	
		/ab_xref="GR:126819"	
		/ab_xref="SWISS-PROT:P54108"	
		/translation="MTPVPLVILVAGLIPSPFANEDKPAFTALITQTOVOREIVIIVH	
		/translation="MTPVPLVILVAGLIPSPFANEDKPAFTALITQTOVOREIVIIVH	
		KHNLRAVSPPRNLWNEAANQKWNQCNTRHSNPKDRMISLKCENLY	
		SSSSNSQAIQSWFEDNFDGIVGPKTFNAYGIVWVWSSLYLGGNAYCPN	
		KVLYVYVQCYCAGNNRNLVYEDGAPCASCPCDNDDGLCINGCKYEDLSNC	
		KLTLKQHLDSCSKNCNSNIV"	
		16..75	
sig_peptide		/gene="CRISP-3"	
BASE COUNT	734	a 397 c 380 g 617 t	
ORIGIN			
Query Match	98.6%	Score 2113.2;	DB 9: Length 2128;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches	2115;	Conservative 0; Mismatches 3; Indels 0; Gaps 0.	
Qy	27	CTGGAACCACTGCAATGCTATTCAGCTGCTGTTCTGGTGTCTGCTGTT	86
Db	1	CTGGAAACCACTGCAATGCTATTCAGCTGCTGTTCTGGTGTCTGCTGTT	60
Qy	87	CCATTTCAGCAATGAGATGCTTACTGCTTACTGCTTACTGCTTACTGCTGTT	146

KEYWORDS	HGCV; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 151752)
TITLE	Homo sapiens, clone RP11-3A4
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 151752)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckley, R., Boguslavsky, L., Boukhalter, B., Brown, A., Castile, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deurrellano, K., Devar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardya, S., Grant, G., Hafos, B., Heaford, A., Hilton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lenoczyk, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, J., Medrini, J., Morrow, J., Naylor, J., Norman, O., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jul 22, 2000 this sequence version replaced g1:7321560.
All repeats were identified using RepeatMasker:	
Smit, A.F.A. & Green, P. (1996-1997)	
http://ftp.genome.washington.edu/RM/RepeatMasker.html	
-----	Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research	
Center code: WIBR	
Web site: http://www-seq.wi.mit.edu	
Contact: sequence_submissions@genome.wi.mit.edu	
-----	Project Information
Center project name: L2999	
Center clone name: 3_A_4	
-----	Summary Statistics
Sequencing vector: M13; M7815; 100% of reads	
Chemistry: Dye-Terminator Big Dye; 100% of reads	
Assembly program: Phrap; version 0.960731	
Consensus quality: 142917 bases at least 040	
Consensus quality: 146940 bases at least 030	
Consensus quality: 148664 bases at least 020	
Insert size: 150752; sum-of-contigs	
Insert size: 150000; agarose-fp	
Quality coverage: 4.8 in 020 bases; agarose-fp	
Quality coverage: 4.8 in 020 bases; sum-of-contigs	
-----	Note: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1 1901: contig of 1901 bp in length	
1 2001: gap of 100 bp	
2 2001: contig of 2630 bp in length	
4632 4731: gap of 100 bp	
4732 4766: contig of 5015 bp in length	
9747 9846: gap of 100 bp	
9847 17668: contig of 7823 bp in length	
17670 17769: gap of 100 bp	
17770 29964: contig of 12195 bp in length	
29965 30064: gap of 100 bp	
30065 43497: contig of 12433 bp in length	
43598 43597: gap of 100 bp	
43598 60340: contig of 16743 bp in length	
60341 60440: gap of 100 bp	
-----	BASE COUNT
ORIGIN	49163 a 27268 c 27224 g 47065 t 1032 others
Query Match	69.3%; Score 1485.2; DB 2; Length 151752;
Best Local Similarity	69.3%; Pred. No. 5e-264; Mismatches 3; Indels 0; Gaps 0;
Matches	1487; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	652 CCAATGGTCAAGTACGAGATCTTACTAAGTTGAAAGTTGAAAGTCACATTAA 711
DB	80106 CCAATGGTCAAGTACGAGATCTTACTAAGTTGAAAGTCACATTAA 80047
QY	712 CCTGTAACATCAGTGGCAGGAGCTGCAAGGCATCTGCATTTGTCACACAGCA 771
DB	80046 CCTGTAACATCAGTGGCAGGAGCTGCAAGGCATCTGCATTTGTCACACAGCA 79987
QY	772 TTATTAATAGCATACACCGACTAGGGTATGTAGAGGGTACAGTATCTAC 831
DB	79986 TTATTAATAGCATACACCGACTAGGGTATGTAGAGGGTACAGTATCTAC 79927
QY	832 TTAGATTTGGCATCTACTAGTATTAACTATACTAGCTGAGAATTGAGCGATTTG 891
DB	79926 TTAGATTTGGCATCTACTAGTATTAACTATACTAGCTGAGAATTGAGCGATTTG 79867
QY	892 ATACACATTTGATTCAGTCAATGTTCTCTGATCTGCTTTTATTACAAATATT 951
DB	79866 ATACACATTTGATTCAGTCAATGTTCTCTGATCTGCTTTTATTACAAATATT 79807
QY	952 TTTCATACAAATGGTAAAGAACAAATCTACAAACACTTGATTATATA 1011
DB	79806 TTTCATACAAATGGTAAAGAACAAATCTACAAACACTTGATTATATA 79747
QY	1012 TAACTTGTGATTAATTAATCTGAATTAGGGAAATTGAAAGTGTATT 1071
DB	79746 TAACTTGTGATTAATTAATCTGAATTAGGGAAATTGAAAGTGTATT 79687

OY	1072	CTCATATGACTAAGTCATAAACCTGGATGAAATTATGTCCTAGAAC	1131	
Db	79686	CTCTATGACTAATGTCATAAACCTGGATGAAATTATGTCCTAGAAC	79627	
OY	1132	AATGTCACAAAGAACATATATTTCACATGAAACCTTGCTGTACTGGCTTCT	1191	
Db	79626	AATGTCACAAAGAACATATATTTCACATGAAACCTTGCTGTACTGGCTTCT	79567	
OY	1192	AGCUCACACTAAGCTTAAAGCTCAAGAGCTTGGCCATTTGCTCTTAC	1251	
Db	79566	AGCUCACACTAAGCTTAAAGCTCAAGAGCTTGGCCATTTGCTCTTAC	79507	
OY	1252	TTCACTCATCATTCACTCTTCACTTCAAGCTTCAAGAGCTTGGCCATTTGCTCTTAC	1311	
Db	79506	TTCACTCATCATTCACTCTTCACTTCAAGAGCTTGGCCATTTGCTCTTAC	79447	
OY	1312	TGTTCTCTTAAACAACTCTGACTTTTGCCTTAACTACATCCCTTGAAT	1371	
Db	79446	TGTTCTCTTAAACAACTCTGACTTTTGCCTTAACTACATCCCTTGAAT	79387	
OY	1372	TTTACATGCTAAATTTCAGAGATGTCACAACCTTCAATTAACTCTCT	1431	
Db	79386	TTTACATGCTAAATTTCAGAGATGTCACAACCTTCAATTAACTCTCT	79327	
OY	1432	CTCTTGACAAACCTCTTAAAGAGATAACAGATAATAGTAAATACCCCTCACT	1491	
Db	79325	CTCTTGACAAACCTCTTAAAGAGATAACAGATAATAGTAAATACCCCTCACT	79267	
OY	1492	CAAGGAGTAGACTCAGTCCTCTCCCTGTGAGCTCTACTAAATCTGACTCACT	1551	
Db	79266	CAAGGAGTAGACTCAGTCCTCTCCCTGTGAGCTCTACTAAATCTGACTCACT	79207	
OY	1552	CCAAAGAGCTGAGATGGAAAGGAAACATAGTAGTACTTCAGGGAGAAAGACAA	1611	
Db	79206	CCAAAGAGCTGAGATGGAAAGGAAACATAGTAGTACTTCAGGGAGAAAGACAA	79147	
OY	1612	TGACTCTCACCAAGTGTCAAATAAAGCTTCACTTCAAGTGTAACTCAGTTG	1671	
Db	79146	TGACTCTCACCAAGTGTCAAATAAAGCTTCACTTCAAGTGTAACTCAGTTG	79087	
OY	1672	CTAGATAATCTCTAAATTCTAAATCCATCTAACTTGTGAGCTAAACATCCAGCA	1731	
Db	79086	CTAGATAATCTCTAAATTCTAAATTCATAATCCCACTAACTATGAGCTAAACATCCAGCA	79027	
OY	1732	AATCAAGTGAAGACATCTACAAATAATCCCTGGGTATTAGTAACTCAGTCAA	1791	
Db	79026	AATCAAGTGAAGACATCTACAAATAATCCCTGGGTATTAGTAACTCAGTCAA	78967	
OY	1792	ACTGAAATACTAGGAAATAGGAACTCTGAGAACATCACAGACCACTGAGACT	1851	
Db	78966	ACTGAAATACTAGGAAATAGGAACTCTGAGAACATCACAGACCACTGAGACT	78907	
FEATURES				
Source				
OY	1852	AGGGACATGTGACCAAATGCAATGGCTCTGGATGAGCTCTGGACAGAAAG	1911	
Db	78906	AGGGACATGTGACCAAATGCAATGGCTCTGGATGAGCTCTGGACAGAAAG	78847	
OY	1912	ATCACTAATCAAACACTGTGAGCTGATGAGTGTGAGTATTAACTGAGTGT	1971	
Db	78846	ATCACTAATCAAACACTGTGAGCTGATGAGTGTGAGTATTAACTGAGTGT	78787	
OY	1972	TTGATTCTTAACTCTGACAAATAGCAGGGTAATGAGTGTAACTCTGAGAAC	2031	
Db	78786	TTGATTCTTAACTCTGACAAATAGCAGGGTAATGAGTGTAACTCTGAGAAC	78727	
OY	2032	TGAACTGGGTGGGGCTATCTGAGATTCTCTGACTACTCTACCAATTTCGTAAG	2091	
Db	78726	TGAACTGGGTGGGGCTATCTGAGATTCTCTGACTACTCTACCAATTTCGTAAG	78667	
OY	2092	TCTAGAACATCCAAATAAAAGTATCTGAAAAAAAGAAAAA 2141		
Db	78666	TCTAGAACATCCAAATAAAAGTATCTGAAAAAAAGAAAAA 78617		

RESULT 6
HSD442L6/C
LOCUS Human DNA sequence from clone RP3-442L6 on chromosome 6. Contains
DEFINITION the TPX1 gene encoding a Testis specific protein 1 (CRISP-2, GADLS), the 5' end of RHKS gene encoding Rhesus blood group-associated glycoprotein (RH50A), the 3' end of SPG28 encoding a cysteine-rich secretory protein-3 (CRISP-3), ESTs, STS and GSS, complete sequence.
ACCESSION AL121950
VERSION AL121950.8
HGNC: GI:9811779
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
Primate; Parker,A.
COMMENT Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CBL15b, UK. E-mail: enquiries@sanger.ac.uk
REFERENCES 1 (bases 1 to 118524)
AUTHORS Parker,A.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CBL15b, UK. E-mail: enquiries@sanger.ac.uk
REMARKS On Jul 28, 2000 this sequence version replaced 91:9408740.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unresolved' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
EmBL; EMBL; SWISSPROT; Tr.; TREMBL; WORMPEP; Information on the WORMPEP database can be found at:
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed from the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at:
<http://www.sanger.ac.uk/HGP/Chr6>
RP3-442L6 is from the library RCR-3 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP3-442L6. The true left end of clone RP3-417L20 is at 85775 in this sequence.
FEATURES
Source
Location/Qualifiers
1. .118224
/organism="Homo sapiens"
/db_xref="taxon: 9606"
/chromosome="6"
/clone="RP3-442L6"
/clone_id="RPCI-3"
/complement("1..25721")
/gene="RHAG"
/gene_type="protein"
/mRNA
/product="RH442L6-1 (Rhesus blood group-associated glycoprotein (RH50A))"
/note="Match: cDNA: Em:X64594 Em:AF179682 Em:AF031548
Em:AF178847 Em:AF178841 Em:AF058917 Em:AF179684
Em:AB01467"
/evidence="not_experimental
complement(join(<1..271279..14163569..37354506..4653,
6950..7100..8061..8244..25538..25694))
/gene="RHAG"
/note="Continues in clone d4417L20 (Em:AL121974)"

```

match: proteins: Tr:Q9V3T3 Tr:088300 Tr:077484 Tr:Q9UBD6
Tr:Q9QXP0"
/codon_start=1
/evidence="not experimental"
/product="DJ472L6.1 (Rhesus blood group-associated
glycoprotein (RH50A))"
/protein_id="CAC10519_1"
/db_xref="GI:1043472"
/translation="MRFTPLMATEVLEIANIVLFLGLFVEVETDQTVLEQINITKPTDM
GIFFLYPIFDVHMFVGFGLMTEFLKKGESVGVINLVLALGLONGTIVGILQ
SQGKFNIGKMINDFSRATVLFQAVLKGTSPTQMLITIEVPHANLYLS
EIKRASDGAISMTTHAFGAVFLAVAGLYNRSGURGHENESEAVSDFAMIGFL
WMPFSNSIAEFGDKCRAIIVYFSLACVLTAFASSSLVHGRKINNMVHIONAT
LAGGVAVGTCADMATHFGSMIGSIAGWVSLGVKFLPFLTTKRL"
repeat_region
/repeat="L1M17 repeat: matches 5986. .6165 of consensus"
repeat_region
/repeat="L1M18 repeat: matches 1. .301 of consensus"
misc_feature
repeat_region
/repeat="match: GSS: Em:AQ813163"
repeat_region
/note="Alusx repeat: matches 3. .300 of consensus"
repeat_region
2602. .2746
/note="L1M3 repeat: matches 7582. .7740 of consensus"
repeat_region
2785. .2832
/note="L1M3 repeat: matches 7478. .7525 of consensus"
repeat_region
3064. .3183
/note="Alusq repeat: matches 1. .311 of consensus"
3823. .4131
/note="Alusx repeat: matches 1. .305 of consensus"
repeat_region
4288. .4406
/note="FLAM_A repeat: matches 1. .122 of consensus"
repeat_region
4820. .5005
/note="MIR repeat: matches 20. .211 of consensus"
5084. .5257
/note="FRAM repeat: matches 0. .172 of consensus"
5951. .6387
/note="L2 repeat: matches 2246. .2698 of consensus"
6673. .6828
/note="MIR repeat: matches 95. .262 of consensus"
complement="7407. .8012"
/feature="RHAG"
repeat_region
/feature="match: GSS: Em:B87311"
repeat_region
/feature="MIR repeat: matches 1. .365 of consensus"
repeat_region
8340. .8442
/note="MIR repeat: matches 4. .106 of consensus"
repeat_region
9022. .9137
/note="MIR repeat: matches 49. .172 of consensus"
repeat_region
9432. .9712
/note="Alusx repeat: matches 1. .282 of consensus"
10074. .10388
/note="L1R16 repeat: matches 14. .381 of consensus"
10815. .10880
/note="33 copies 2 mer tt 66% conserved"
10939. .11098
/note="THELB repeat: matches 1. .167 of consensus"
repeat_region
11099. .11389
/note="Alusq repeat: matches 2. .297 of consensus"
repeat_region
11390. .11609
/note="THELB repeat: matches 167. .364 of consensus"
repeat_region
11713. .12019
/note="Alus repeat: matches 1. .304 of consensus"
repeat_region
12444. .12662
/note="MER0 repeat: matches 1. .217 of consensus"
repeat_region
13096. .13085
/note="Alusx repeat: matches 2. .290 of consensus"
repeat_region
13474. .13533
/note="MIR repeat: matches 74. .182 of consensus"
complement="13493. .13920"
/feature="RHAG"
/feature="match: GSS: Em:AO813968"
repeat_region
13541. .13606

```

```

misc_feature
/repeat="L2 repeat: matches 2637. .2707 of consensus"
13971. .14526
/note="match: GSS: Em:AO539951"
repeat_region
14139. .14391
/note="MIR repeat: matches 5. .262 of consensus"
repeat_region
14620. .14983
/note="THELB repeat: matches 1. .364 of consensus"
repeat_region
14984. .16480
/note="THELB-INTERNAL repeat: matches 1. .1580 of
consensus"
repeat_region
16481. .16840
/note="THELB repeat: matches 1. .364 of consensus"
repeat_region
17057. .17296
/note="Alus repeat: matches 31. .282 of consensus"
repeat_region
17297. .18994
/note="L1M5 repeat: matches 743. .1686 of consensus"
repeat_region
19639. .19831
/note="L1M3 repeat: matches 7448. .7634 of consensus"
repeat_region
19863. .20630
/note="L1M4 repeat: matches 6630. .7474 of consensus"
repeat_region
20891. .21016
/note="Charliela repeat: matches 1. .128 of consensus"
repeat_region
21013. .21609
/note="Charliela repeat: matches 331. .933 of consensus"
repeat_region
22116. .22212
/note="L2 repeat: matches 2589. .2695 of consensus"
repeat_region
22790. .22844
/note="L2 repeat: matches 2617. .2671 of consensus"
repeat_region
22860. .23218
/note="L1M4 repeat: matches 5950. .6299 of consensus"
repeat_region
23216. .23511
/note="L1M2 repeat: matches 5508. .5806 of consensus"
repeat_region
23683. .23632
/note="L2 repeat: matches 6152. .6302 of consensus"
repeat_region
23839. .23980
/note="21 copies 2 mer tg 88% conserved"
repeat_region
24216. .24538
/note="match: GSS: Em:AO701395"
repeat_region
27538. .27945
/note="L1M9 repeat: matches 2356. .2750 of consensus"
repeat_region
28904. .29192
/note="Alus repeat: matches 1. .289 of consensus"
repeat_region
29656. .30138
/note="Charlie4a repeat: matches 1. .507 of consensus"
repeat_region
30286. .30707
/note="match: GSS: Em:AO465695"
repeat_region
30451. .30583
/note="MIR repeat: matches 1. .134 of consensus"
misc_feature
30513. .30810
/note="match: GSS: Em:AO114043"
repeat_region
30770. .31070
/note="Alus repeat: matches 1. .303 of consensus"
repeat_region
31388. .31559
/note="36 copies 2 mer ta 69% conserved"
repeat_region
31808. .32127
/note="Alus repeat: matches 1. .309 of consensus"
repeat_region
32391. .32514
Query Match 69.0%; Score 1480.4; DB 9; Length 118524;
Best Local Similarity 99.6%; Pred. No. 4e263; Mismatches 6; Indels 0; Gaps 0;
Matches 1484; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 652 CCGATGGTGCAGTACGGAGATCTTATGACTGTTGAGCTCACATTA 711
Db 117739 CCAATGGTGCAGTACGGAGATCTTATGACTGTTGAGCTCACATTA 117680
QY 712 CCTGTAACATCAGTGGTCAGGACAGTCAGCTCTGCAATTGTCACAGCA 771
Db 117679 CCTGTAACATCAGTGGTCAGGACAGTCAGCTCTGCAATTGTCACAGCA 117620
QY 772 TTATTAATACCGATACACCCGAGTAGGGCTATGAGAGGAGTCAGATTCTAC 831
Db 117619 TTATTAATACCGATACACCCGAGTAGGGCTATGAGAGGAGTCAGATTCTAC 117560

```

Qy 832 TTAGATGGCATCTACTTAGATTAACATATCAGTGAGAATGTAGGCATGTT 891
 Db 117559 TTAGATGGCATCTACTTAGATTAACATATCAGTGAGAATGTAGGCATGTT 117500
 Qy 892 ATACACATGTTCAATGTTTCTCGATCTGCTTTAGTACATTAACATCTACTAGCTGAGAAATGTAGGCATGTT 951
 Db 117499 ATACACATGTTCAATGTTTCTCGATCTGCTTTAGTACATTAACATCTACTAGCTGAGAAATGTAGGCATGTT 117440
 Qy 952 TTTCATACAAATGGTTAAAAGAACAAATCTATACACACACTTGGATTATATA 1011
 Db 117439 TTTCATACAAATGGTTAAAAGAACAAATCTATACACACACTTGGATTATATA 117380
 Qy 1012 TAACTTGTGTTAAATTCTAGTGGATTATGGTGAAGATTGGTGAAGTTGAGTGATT 1071
 Db 117379 TAACTTGTGTTAAATTCTAGTGGATTATGGTGAAGATTGGTGAAGTTGAGTGATT 117320
 Qy 1072 CTCATATGACTAGTCACAAACCTGGATGAAATATGTCTAGACA 1131
 Db 117319 CTCATATGACTAGTCACAAACCTGGATGAAATATGTCTAGACA 117260
 Qy 1132 ATATGTCACAAAGAACAAATTAATTTCAGATGAAACCTGGCTTTCT 1191
 Db 117259 ATATGTCACAAAGAACAAATTAATTTCAGATGAAACCTGGCTTTCT 117200
 Qy 1192 AGCTCCACTCTAAGGCTAAGCATCTTCAGTCAGTGAACTCTGGCTTAGTGGCTTCT 1251
 Db 117199 AGCTCCACTCTAAGGCTAAGCATCTTCAGTCAGTGAACTCTGGCTTAGTGGCTTCT 117140
 Qy 1252 TTCACTCATTGACCCCTCTCCAACTCATCTGCGTCATCTCAGTGAAAGCTTCCGATCTGCTTAATCT 1311
 Db 117139 TTCACTCATTGACCCCTCTCCGATCTGCTTAATCT 117080
 Qy 1312 TTTCCTCTCAAACATCTGACTTTGCAAAATCAATCAATCCCTTGATT 1371
 Db 117079 TTTCCTCTCAAACATCTGACTTTGCAAAATCAATCAATCCCTTGATT 117020
 Qy 1372 TTTCATCTGATAATTAGTAGATAATGTCACCCCTCATTTAACCTCTT 1431
 Db 117019 TTTCATCTGATAATTAGTAGATAATGTCACCCCTCATTTAACCTCTT 116960
 Qy 1432 CTCTTGACAAACTCTCTAAAGAACATAATATGCTAACCTCCACT 1491
 Db 116959 CTCTTGACAAACTCTCTAAAGAACATAATATGCTAACCTCCACT 116900
 Qy 1492 CAAGAGGCTGACTCTCTCCCTGAGCTCTCAATAATCAGTGCCTACTT 1551
 Db 116899 CAAGAGGCTGACTCTCTCCCTGAGCTCTCAATAATCAGTGCCTACTT 116840
 Qy 1552 CCAAGAGTGGAGTATGAAAGGGAAACATAGTAACCTTACGGGAGAAATGACAA 1611
 Db 116839 CCAAGAGTGGAGTATGAAAGGGAAACATAGTAACCTTACGGGAGAAATGACAA 116780
 Qy 1612 TGACGTTCTACCAAGTGTCAAAATAAGTCACCGTGTAGTCATTGAGTTGTT 1671
 Db 116779 TGACGTTCTACCAAGTGTCAAAATAAGTCACCGTGTAGTCATTGAGTTGTT 116720
 Qy 1672 CTAGATAATCTCTAAATTCTAAATCCATTAATGCTAAACATGACCGCA 1731
 Db 116719 CTAGATAATCTCTAAATTCTAAATCCATTAATGCTAAACATGACCGCA 116660
 Qy 1732 ACTCAAGTGGAGGACATCTACAAATTCCTGGGTATTTAGAGTATCTCAA 1791
 Db 116659 ACTCAAGTGGAGGACATCTACAAATTCCTGGGTATTTAGAGTATCTCAA 116600
 Qy 1792 ACTGTTAAATCTGGAATAATSGGAATCTCTGAGAACATCACAGACCATGAGCT 1851
 Db 116599 ACTGTTAAATCTGGAATAATGGGAATCTCTGAGAACATCACAGACCATGAGCT 116540
 Qy 1852 AAGGAGACATGTCGCCAATGCAATGCTCTGGATCAGATCTGGACAGAGAAAG 1911
 Db 116539 AAGGAGACATGTCGCCAATGCAATGCTCTGGATCAGATCTGGACAGAGAAAG 116480

RESULT 7
 HUMTPXIA HUMTPXIA 1335 bp mRNA linear PRI 14-JAN-1995
 LOCUS HUMTPXIA
 DEFINITION Human testis-specific protein (Tpz-1) mRNA, complete cds.
 ACCESSION M25532.1 GI:339882
 VERSION M25532.1
 DEFINITION testis-specific protein.
 SOURCE Human adult testis, cDNA, clones H4-1 and p3-1.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Buteleostomi;
 Kasahara,M., Gutknecht,J., Brew,K., Spurr,N. and Goodfellow,P.N.
 AUTHORS Cloning and mapping of a testis-specific gene with sequence
 TITLE similarity to a sperm-coating glycoprotein gene
 JOURNAL Genomics 5 (3), 527-534 (1989)
 MEDLINE 90129048
 PUBMED 2613236
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
 by M.Kasahara, 09 JUN-1989.
 FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="6p21-qter"
 gene
 1..1335
 /gene="GAPDL5"
 mRNA
 <1..1335
 /gene="GAPDL5"
 CDS
 /product="testis-specific protein mRNA"
 190..921
 /gene="GAPDL5"
 /note="testis-specific protein precursor"
 /codon_start=1
 /protein_id="AA6120.1"
 /db_xref="GI:339883"
 /db_xref="GDB:00-119-257"
 /translation="MALLPFLVPLSLPAGKDDPAFTALITLQLQVQREIVKHNELIKAVSPASNLKMEVNSREVNIAOWANCKTLOSDPERKTSRPGENLYMSSDPTSSRATOSWDEILDVYGVPSKPVAVVGHITQWYSTIQVGGIATCPNPDLSKLYVYVQCCPAGNNMRKTPYQGTCAGCPDCDDGKLTNSQYDPLSNCDSLNNTACCEHLLKEKKATCLENKIV"
 190..249
 /gene="GAPDL5"
 /note="testis-specific signal peptide (put.); putative"
 mat_peptide
 /gene="GAPDL5"
 /product="testis-specific protein (put.); putative"
 BASE COUNT 428 a 279 c 271 g 355 t 2 others
 ORIGIN Chromosome 6.
 Query Match 24.0%; Score 515.2; DB 9; Length 1335;
 Best Local Similarity 68.8%; Pred. No. 4.4e-85;
 Matches 798; Conservative 0; Mismatches 323; Indels 39; Gaps 5;

Qy 1 TGATGAAACAAATCTCACTCTCTGCTGAAACCA---CTGCAATGCAATTATCCCCA 56

Db	142 TATAAAGTAGTATTCATCCTGCCTAGAAACATTTCCACCAATGGTGCATTACGTATTCTGAAACAT 201	Db	1188 TAGGATTAGGTTCACTTAGACTTGGATCAAATGGTGCATTACGTATTCTGAAACAT 1247
Oy	57 GTGCTGTGTTCTGTTGCTGGGCTGCTTCCATTTCCAGCAATGAGATAGGAT 116	Oy	1136 GTCACAAAGAACAAATATAA 1155
Db	202 CGGCTGTGTTCTGTTACTGTGCTGCTTCACTTTACGTAA 258	Db	1248 GCTAACAGAGACTGTAA 1267
Oy	117 CCCGCTTTACGTCTTGTAAACCACCCAAACACAAGTCGCAAGGGAGATGTGAAATAG 176		RESULT 8
Db	259 CCCGCTTTACGTCTTGTAAACCACCCAGTGCAGNTGCAGGGAGATGTGAAATAG 318	BC022011	BC02011 1380 bp mRNA linear PRI 24-JAN-2002
Oy	177 CAAATGAACTGGAGGAGCAGTATCTCCCTGCCAGAAACATGGTGAAGATGGATGG 236	DEFINITION	Homo sapiens, testis specific protein 1 (probe H4-1 p3-1), clone MGC:26358 IMAGE:4826427, mRNA, complete cds.
Db	319 CAAATGAACTGGAGGAGCAGTCTCCACCTGCCAGTAAAGCTGAAAGATGGATGG 378	ACCESSION	BC02011
Oy	237 AACAAAGAGGCTGCAGCAATGCCAAAGTGGCAACACAGTCGCAATRCAGACAGT 296	VERSION	BC02011.1 GI:18314472
Db	379 AGCAGAGGATPACACGAAAGTGGCAACAAAGTGGCAACAACTGACTTACACATAGT 438	KEYWORDS	MGC.
Oy	297 AACCCAAAGGATCGAATGACAGTCATAAAATGGTGGAGAACTCTCATATGCACTG 356	SOURCE	Homo sapiens.
Db	439 GATCCAGAGGACGCCAAACCGTACAGAGTGGTGGAGAACTCTCATATGCACTG 498	ORGANISM	Homo sapiens.
Oy	357 CCCAGCTCATGTCAPACAGCAATGCCAAAGTGGCAACAACTGACTTACACATAGT 416	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. (bases 1 to 1380)
Db	499 CCTACTCTGTCTCTGCATCCAAAGCTGCTGATGACGAGATCTTAGTTGACTT 558	AUTHORS	1. Strausberg, R.
Oy	417 GGTGTTAGGGCCAAAGTCCCACCGACTGGTGGACATTACAGGTGTTGGTAC 476	TITLE	Direct Submission
Db	559 GGTGTTAGGGCCAAAGTCCCACCGACTGGTGGACATTACAGGTGTTGGTAC 618	JOURNAL	Submitted (22-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
Oy	477 TCTTCATACCTGTTGATGTTGGAATGCCTACTGTCCTAACTAAAGTCCTAAATAC 536	REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
Db	619 TCGACTTACCAAGTAGCTGTTGAATGCTCCAACTCAAACTGCTAAATAC 678	COMMENT	Contact: MGC help desk Email: cgapbs@rmmail.nih.gov
Oy	537 TACTATGTTGCCAATATGTCCTGCTGGTAATGGCTTAATAGCTATATGCCCITAT 596	Tissue	Procurement: Miklos Palkovits, M.D., Ph.D.
Db	679 TACTATGTTGCCAATATGTCCTGCTGGTAATATGAAAGATACCGCTAC 738	Preparation:	CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Yoshiyuki and Piero Carnini (RIKEN)
Oy	597 GACAAAGGAGCCTGTGCCAGTACTGTGAGATGACTGACCAAT 656	Series:	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Db	739 GACAAAGGAGCCTGTGCCAGTACTGTGAGATGACTGACCAAT 798	DNA Sequencing by:	Institute for Systems Biology http://www.systemsbiology.org
Oy	657 GCTTCAAGTAGCAGAAGTCTCTATGAACTCTAAAGTTGCAAGCTCACATAACCTGT 716	contact:	amadane@systemsbiology.org
Db	799 ACTTGGCACTATCAAGATCTCTTAAGTAACTAGCTGGCT 858	Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting	
Oy	717 AACATCAGTGTGTCAGGGACATCTGCAATTTGTCAAACGCCATTA- 775	FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium LNL at: http://image.lnl.gov
Db	859 GACATGAGTTACTCAAGGAAGACTGCAAGGCTACTGTGCTATGTGAGACCAATTAC 918	source	Series: IMAK Plate: 33 Row: n Column: 2 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507670.
Oy	776 TAAATGGCACTACACCGGTAGGGCTATGAGAGGGTCAAGTACTACCTAG 835	location/Qualifiers	1. .1380.
Db	919 TCAATTACCTACTGAGGATGTGCAAGACTGGATGGATAAGGCTGTGATCTTAATT- 976		/organism="Homo sapiens" /db_xref="LocustID:7180" /db_xref="Taxon:9606" /clone="MGC:26358 IMAGE:4826427" /tissue_type="Testis" /clone_id="NIH-MGC_97" /lab_host="DH10B" /note="vector: PBluescript" /vector="PBluescript" /note="vector: PBluescript" /vector="PBluescript"
Oy	836 ATTTGGCATCTACTTACATTACATACTGCTGAAATGTGAGCATGTTGATAC 895	CDS	219. .950
Db	977GCGACATACAGGGAAATGTGATGTGATGTGAC 1013		/codon_start=1 /product="testis specific protein 1 (probe H4-1 p3-1)" /protein_id="AAH22011" /db_xref="GI: 18314473"
Oy	896 ATTTGGCATCTACTTACATTACATACTGCTGAAATGTGAGCATGTTGATAC 955		/translation="MALLPVFLVILPLSLPAEGKDPAFFLTLTQLOVOREYVNH NELRKAVSPSPNMLKNEWSREYVNTQWRANKTQHSDPEKRISSTRGENTLNS DPTSWSSAISQYDEIDFVQVGPKSPNAYGHYQOLWVISTYQVQGJAYCPNODS LKVYVQYCPAGNNMRKPNQYQCPGCPDDKGGLCTNSCOYQDLSNCOSL NTAGCEEDLKEKCATCLCENRY"
Oy	1014 ATTTGGCATCTACTTACATTACATACTGCTGAAATGTGAGCATGTTGATAC 1073	BASE COUNT	448 a 289 c 280 g 363 t
Db	1132TTAGTGTGTTATACGACTGAACTGAGATTTGAGATTTCTA 1187	ORIGIN	Query Match: 24.0%; Score 513.6; DB: 9; Length 1380; Best Local Similarity: 68.7%; Pred. No.: 8.6e-85; Matches: 777; Conservative: 0; Mismatches: 324; Indels: 39; Gaps: 5;
Oy	1076 TATGACTAAGTCACTAAACCCGGATGAAATGTGCTCTAGAACAAAT 1135	Qy	1 TGTGAAACAAATACTTCATCCGCTCGAACCA---CTCAATGACATPATCCCA 56

		VERSION	AJ001400.1
		KEYWORDS	GI:2388784
		SOURCE	CRISP-3; cysteine-rich secretory protein-3.
		ORGANISM	Equus caballus
Db	370	CAACATGAACTAGGAAGGAGCTCTCCACCTGCCAGTAACATGCTAAGATGGATGG	429
Oy	237	ACAAAGAGGCTGCAGCAATTGCCAAAATGGGCCAACCGAGTCATACAGACACT	296
Db	430	AGCAGAGGAGTACACRAGGATGCCAACGGCAGAACAGTCACTTACACATAGT	489
Oy	297	ACCCAAAGGTCGAATGACAGTCATAAATGGTGGAGAACTCTACATGCTAAGTGCC	356
Db	490	GATCCAGAGGACGCCAAACCGTCAGAAGTGGTGGAGAACTCTACATGCTAAGTGAC	549
Oy	357	CCAGCCTATGGTCACAAGCACTCCAAAGCTGGTGTGATGAGTACATGATTGACTT	416
Db	550	CTTACTCTCTGCTCTGCTACCAAGCTGATGACAGGATCTAGATTTGCTAT	609
Oy	417	GGTGTAGGGCCAAAGACTCCACCGAGTGGTGGACATATACAGGTTGGTGTAC	476
Db	610	GGTGTAGGACCAAGAGTCCTACATGCTGAGTGGTGTGACATATACAGCTGGTAC	669
Oy	477	TCTCATACCTCTGGATGGTGGAAATGCCCTACTGCCAACATAAAAGTCTAAATAC	536
Db	670	TGACGTTACCAAGTAGGCTGTTGAGAATGCTCTACATCAAGATASTCTAAATAC	729
Oy	537	TACTATGTTGCCAATATGCTCTGGTAAATGGGCTAATAGACTATGTCCTAT	596
Db	730	TACTATGTTGCCAATATGCTCTGGTAAATATGATAGAAGATACTCCCGTAC	789
Oy	597	GAACAAAGGAGCCTGTGCGGTGCGCTGAGTACTGTGACAAAGACTATGACCAAT	656
Db	790	CAACAGGAAACCTGTGCGGTGCGCTGAGTACTGTGACAAAGACTATGACCAAT	849
Oy	657	GGTGCAGTACAGTACGAACTCTATGTAACGTGAACTGTGACATTAACCTG	716
Db	850	AGTGGCAGTACAGTACAGTCTCTAAGTAACTGTGATCTGAGAATACAGCTGGT	909
Oy	717	AAACATCAGTGTGTCAGGGAGCTGGCAAGGCACTCTGCAATGTCAAAGACATTA	775
Db	910	GAACGAGGTGTCTCAGGAAGAAGTGCAGAAGCTACTGCTCCTATGAGAACATTAC	969
Oy	776	TTAAATAGGCATACACCGAGTAGGGCTAGTGGAGAGTCAGATTACTTAG	835
Db	970	TGATTCCTAGTGGAGCATGTGCAAGACTGCATGGATAAGGCTGCATTTAATT-	1027
Oy	836	ATTGGCATCTACTTAGTTAACATACTAGCTGAGAAATGGTGGCAGTTGATAC	895
Db	1028	-----GGCACATACCACTGGTGAATGTTGATGATGTTAGTGTAC	1064
Oy	896	ACATTGATTCAATGTTCTCTGGATCTGCTTITTTCAAAATATTT	955
Db	1065	AAATTGATTCAACAGCAATGCTCTCCCGATCATCACAGAAATCTCTCA	1124
Oy	956	ATACAAATGTTAAAGAACAACTTACACAACTTGGATTATATATAAA	1015
Db	1125	GGCANTGATTCAAAAGTGGCATGATGATGATGACACTGTGACTCTGCAAA-	1182
Oy	1016	CTTGTGATTAAATTACTGATTAATTGGTAAATTGTAAGTGTATCTCA	1075
Db	1183	---TTAGCTTAAACGACTGATAGTGGATTGAAACTGTATACCA	1238
Oy	1076	TATGACTAATCTACTAAACCTGGATTGAAAGGAAATATGTCCTGAGAACAAAT	1135
Db	1239	TTGGATTAGGACTGGATCAAAATGGTCACTGTATTCTGAAACAT	1298
Oy	1136	GTACAAAGAACAAATAA	1155
Db	1299	GCTAAGAAGAACTGTA	1318
RESULT	10		
ECRISP3		1295 bp mRNA	linear
LCUS		Equis caballus	MA03-FEB-1998
DEFINITION		Equis caballus mRNA for cysteine-rich secretory protein-3.	
ACCESSION		AJ001400	
FEATURES			
source		Location/Qualifiers	
	1. .1295	/organism="Equus caballus"	
		/db_xref="taxon:9795"	
		/sex="Male"	
		/tissue_type="ampulla"	
		/dev_stage="adult"	
	59. .796	/gene="crisp-3"	
		/gene="crisp-3"	
		/codon_start=1	
		/product="cysteine-rich secretory protein-3"	
		/protein_id="CAA04729.1"	
		/db_xref="GI:2388785"	
		/db_xref="SWISS-PROT:019010"	
		/translation="MALLVPLFLAWLIPPFPPASQDGFALSLITKSEVOKEIVNK	
		HNDLRTIVSPIASNMKMQNDSKTNAQWANKLILQHKAEDRAGTMKGEMLFM	
		SSPNWSQDANWIDEVHPKYGPKTPNAWPHYTQWWSYRVCQGPKQ	
		CTKVVYVCOCPNQVNNVYVQGPKQ	
		LKKIACGEHELKENKENCYTCOCENKY"	
	59. .124	/gene="crisp-3"	
		polyA_signal	
	1262..1267	1262..1267	
ORIGIN		BASE COUNT	
	428 a	257 c	
		245 g	
		365 t	
Query Match		Best Local Similarity	23.6%
		Score	505.8
		DB	4
Matches		Pred.	2.4e-83
	803; Conservative	Mismatches	0;
		Indels	297;
		Gaps	39;
Oy	4	TCAACAAATCTCATCTGCTCTGCAACACTGCAATGACATTATGCCAGTCGT	63
Db	24	TAAAGCAGATGCTCTCACCTTGCTGCTGAGAACCA--ACATGCCATTATCAGCTGTC	80
Oy	64	TGTCCTGTTCTGGCTCTCCATTTTCAGCAATGAGATAGGATCCCTT	123
Db	81	TGTTTCGCTGCCGCTGCTGCTCACCTTCTGCAAGTGA--CAGATCCAGTT	137
Oy	124	TRACTGCTTGTAAACCCAAACAGTCAAAGGAGATGAGATAGCACATA	183
Db	138	TGCTGTTGTCATCACCAAAAGGCAAGTCCAAAGAGATGTAATACACATG	197
Oy	184	AACTGAGGAGGAGCTGCTCTCCCGCCAGAACAGTCATGAGATGGAACAAAG	243
Db	198	ACCTAAGGAGAACAGTCCTCCACTTCAGTCATGCTAAAGATGCGACAGCA	257
Oy	244	AGGCTGAGCAATGCCAAAGTGGCCAAACAGTCATGCAAGTACAGTACCA	303
Db	258	AGACAGCAAAATGCCAAACTGGCCAAAGTGGCTCTCCACACAGTAAAGCAG	317
Oy	304	AGGATCG--AATGACAGTCATAATGGTGGAGATCTCATGTCAGTGC	360

	MEDLINE	96354287
REFERENCE	PUBMED	9115720
	2 (bases 1 to 1388)	
AUTHORS	FOSTER, J. A. and GERTON, G. L.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-SEP-1995) James A. Foster, Department of Obstetrics and Gynecology, Division of Reproductive Biology, University of Pennsylvania, 36th and Hamilton Walk, Philadelphia, PA 19104, USA	
FEATURES	source	1. 1388
CDS	/db_xref="GI:1016756" /cell_type="sperm acrosome" /tissue_lib="testis" /cood_start=1 /product="autoantigen 1" /protein_id="AAC52616.1" /db_xref="GI:1016756" /translation="MALLPVVFLITMLLPCVLTNCDPARTALITQSQVNELINKHNRQKSYPPASNNPMLMREANVNAQCNPMLRQEVLPVNPDRKSTCQGENLUMSDPSSWDATQSFDEQSDTFVGPKSHAVVHGHTQWLYISLVGGIACRNQDSLKYYVCOCPAGNNVYKNTPYKQGIPCAASCOPGHENGICNTNSEYEDLISNECSL	
misc_feature	/note="encodes region previously determined by microsequencing (Hardy et al. Biol Reprod vol. 38, p. 423-437, 1988; PIR Accession Number A68022, no/	
BASE COUNT	444 a 276 c 285 g 383 t	
ORIGIN		
Query Match	22.2%	Score 476.2; DB 10; Length 1388;
Best Local Similarity	68.2%	Pred. No. 6. 5e-8;
Matches	758;	Mismatches 303; Indels 40; Gaps 6;
QY	841	GCATCTACTTGTAGTTAACATACTAGCTGAGAAATTAGCATGTGATACATT 900
QY	835	GCTGGATCCCTTGTGACATACCCG - AGGAANTTATGGCATGCTGTGCAAGC 893
QY	901	TGATTCAAAGTTCTCTGGATCTGCTGTTTATTACAAANTATTTCATACA 960
QY	894	TGATCCAA - CCGTAATGCATTTCTGTACATTACAGAATCTCTCCACACA 951
QY	961	AATGGTTAAAGAACAAACTTACACACAACTTGGAATTATATAACCTG 1020
QY	952	ATGATTACAAAGCAGTACTCTGTGATGACAATTGTGCTT - GATAATTG 1008
QY	1021	TGATTAATCTACGATTAATAGGGAAATTGGAAGTGTATTCATG 1080
QY	1009	TACTTTAATGTAATTAATTGATCA - ATGGAGATTGTGAAAGTGA 1067
QY	1081	CTAAGTCACTAAACCTGGATGAAAGTGAATTATGTCTCTAGAC 1139
QY	1068	CTTAGTCCTGAACTTGTGATTAACAGGAATTAACGATCTCTGAAACACATGC 1126
RESULT 1		
CP035712	CP035712	1388 bp mRNA linear ROD 18-JUN-1996
LOCUS	Cavia porcellus	1388 bp mRNA linear ROD 18-JUN-1996
DEFINITION	Cavia porcellus acrosomal autoantigen 1 mRNA, complete cds.	1388 bp mRNA linear ROD 18-JUN-1996
ACCESSION	U35712	1388 bp mRNA linear ROD 18-JUN-1996
VERSION	U35712.1	1388 bp mRNA linear ROD 18-JUN-1996
KEYWORDS		
KEYWORDS		
SOURCE		
ORGANISM	Cavia porcellus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystericognathi; Caviidae; Cavia. (bases 1 to 1388)		
REFERENCE	Foster, J. A. and Gerton, G. L.	
AUTHORS		
TITLE	Autoantigen 1 of the guinea pig sperm acrosome is the homologue of mouse Tpx-1 and human Tpx1 and is a member of the cysteine-rich secretory protein (CRISP) family. Mol. Reprod. Dev. 44 (2), 221-229 (1996)	
JOURNAL		

Db	674	TAATCTACTAGTGTGCCAATACTGTGCCCTGGTAAATAGTGTATACAAAGAACACTCT	733
Qy	594	TATGAAAGAGGACACCTTGTGCCAGATAACTGTGACGATGACTATGACCC	653
Db	734	TATAACAAAGGAACTCTTGCCAGCTGCCCTGACATGTGAAATGGACTATGACCC	793
Qy	654	AATGGTGTGCAACTAGGAGATCTATGTAAGTAAAGTGTGAGCTCATTAAAC	713
Db	794	ACAGTGTGAGTGTGAGATCTCTAGTAAGTGTGAGCTTGTGAGATACAGCTGGC	853
Qy	714	TGTAACATCAGCTGGCAGGCACTGTGCAAGCTCCTGCAATGTCACACGCACT	773
Db	854	TGTAACATCAGCTGGCAGGCACTGTGCAAGCTCCTGCAATGTCACACGCACT	913
Qy	774	TATTAATACGCTTACCAACCCAGTAGGGCTGTAGAGGAGTCAGATATCTACT	833
Db	914	TACTAAACCTACGAGCTGTGAGGACTA-----GAGGAGGAGGAGGAGGAG	946
Qy	834	AGATGGCCTACTAGATTACATATACAGCTGAGAAATTGTGAGCATCTGT	893
Db	947	AAACAGGGCACCATTGGTGGACACAAACCAAGATAAACCTAGTAGTGTAGT	1006
Qy	894	ACACATTTGTTCAAAATGTTCTCTGGATCTGCTTTATTTACAAATATT	953
Db	1007	ATAGTTAGATGCCAATAGGACTACTACATAGTGTCTGGATTTAGAAATTGT	1066
Qy	954	TCATACAAATGGTTAAAGAACAAATCTATACACAACTTGGATTATATA	1013
Db	1067	TGTCAGCAGCAATTACAAAGAACAGCTGTCT-ATAACATCTGGACTTGATAGC	1125
Qy	1014	AACCTGGGATTACTGATTATAGGGTAAATTGTGAAAGTGTGATT	1073
Db	1126	A-TTCATCTTAAATTATGATTGATCAGATCAGATCAGTGTGACT	1183
Qy	1074	CATATGACTAAGTCATAAA 1094	
Db	1184	CATAGGACTTAGGTCACIAGA 1204	
RESULT 12			
MUSTPXA	MUSTPXA	1418 bp	mRNA
DEFINITION	Mouse testis-specific protein (Mpx-1) mRNA, complete cds.		linear
ACCESSION	M25533	J07471	ROD
VERSION	M25533.1	61:202126	27-APR-1993
KEYWORDS	testis-specific protein.		
SOURCE	Mouse (strain Balb/c) testis, cDNA to mRNA, clone B4(1)4.		
ORGANISM	Mus musculus		
BIOLOGICAL_ASSEMBLY	Bukarrotta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1418)		
AUTHORS	Kasahara,M., Guttmrech,J., Brew,R., Spurr,N. and Goodfellow,P.N.		
TITLE	Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene		
JOURNAL	Genomics 5 (3), 527-534 (1989)		
MEDLINE	90129048		
PUBMED	2613235		
COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted by M. Kasahara, 09-JUN-1989.		
FEATURES	Location/Qualifiers		
source	1. .1418		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
mRNA	<1. .1418		
	/product="testis-specific protein mRNA"		
CDS	214. .945		
	/note="testis-specific protein precursor"		
	/codon_start=1		
	/protein_id="AA44972.1"		
	/db_xref="GI:202126"		
	/transl_table="MAFWQMLFVALLRLSPLEPGKDPDFTSLTNOLOVORELVNK		
	HNELRSSNPISDILKNEWSIQATINQWANKCILSERHSSKDRKINRCGENIYS		

Db	1116 TTGTGTCTAACCAACCTGCAATGGAAAGAA 1146	Db	-416 ACATACTAAATGGAACTGACAGCAGCACCAATGCTCAAAGTGGGTAATA 475
RESULT	14	QY	277 AGTCATAGAGACAGTAACTCCAAAGGTGAAAGCAGCTAAATGTCGAGA 336
DEFINITION	AF078552 AF078552 Rattus norvegicus cysteine-rich secretory protein-2 CRISP-2 (tpx-1) mRNA, complete cds.	Db	476 ACTGATTTAGAACACAGTGTACAGAGAACGGAAATCAATTAATGTCGGAGA 535
ACCESSION	AF078552 AF078552	QY	337 ATCTCTACATGTCAGTGCCGCCAGCTATGGTCAAGCAATCCAAAGCAGCTGGTTGATG 396
VERSION	AF078552.1 GI:5725525	Db	536 ATCTCTATGTCAGTGACGCCCTACATCTCTGAGAACGGTATTCAGCTGGATGAG 595
KEYWORDS	Rattus norvegicus.	QY	397 AGTACAAATGATTTGACTTGTGAGGCCAAAGCTCCACGCGATGTTGACATT 456
SOURCE	Rattus norvegicus.	Db	596 AAAATGAAACTCTCGTTTCGCGCTAGGCTAA---ACCCAAATCGCTGCGGACAT 652
ORGANISM	Rattus norvegicus.	QY	457 ATACACAGGTTCTTGTACTCTCTCATCTCTGTTGATGAGGAATGCCACTGTCCA 516
JOURNAL	Mol. Reprod. Dev. 50 (31), 313-322 (1998)	Db	653 ACACTCAGCTGTTGGTATCCTTCAGTGTGGAGSTGCTFACTGRCCA 712
MEDLINE	98294327	QY	577 ATAGACTATATGCCCTATGACAAAGGACCTGTCGAGTGGCCAGATAACTGTG 636
REFERENCE	1 (bases 1 to 1432)	Db	517 ATCAAAGATCTAAATACTACTATGTTGCCAATATGTCCTGCTGTTATTGGCTA 576
AUTHORS	O'Bryan, M.K., Loveland, K.L., Herszfeld, D., McFarlane, J.R., Hearn, M.T. and de Kretser, D.M.	QY	773 TGAAGAAAGATGACCCCATATCATCAGGACACCTGCTGCTAGTGGCTATGGTACACAGTGA 772
TITLE	Identification of a rat testis-specific gene encoding a potential rat outer dense fibre protein	Db	713 ATCAAGAATACCCCTGAATACTCTCTATGTTGCCATTACTGCTCATAAGTT 696
JOURNAL	Direct Submission	QY	637 AGGATGGACTATGCCACCAATGTTGCAAGTAGAATCTCTATAGTT 697
MONASH	Submitted (15-JUL-1998) Institute of Reproduction & Development, Monash University, Monash Medical Centre, 246 Clayton Rd., Clayton, Victoria 3168, Australia	Db	833 ATAATGCGATTGTCGACCAATGCTGTTGAGAATCTCTATAGTTAACAGTT 892
FEATURES	Location/Qualifiers	QY	697 TGAGCTCACAATTAACCTGAAATCATGTTGGTCAGGACAGTTGCAAGGATCTGCA 756
source	I.. 1432	Db	893 TGAGAGTTCAGCACCGCTGTAACTGATGTTCTCANGCAAGTGTGAGGCTACTGCC 952
organism	"Rattus norvegicus"	QY	757 ATGGTCAACACGATTTAA-ATAGCATTACACCCGAGTAGGGCTATGAGAG 815
strain	"Sprague-Dawley"	Db	953 TATGTGAGCACAAATTCATTACATGCCAGCGTGCAGCATGAGGG 1012
gene	I.. 1432	QY	816 GACTCAGATTACTACTAGTTGGCTACTTAGTTACATAGTTAACATATAGCTGAGAA 875
CDS	/gene="tpx-1"	Db	1013 GGGTACAGACTTAGTGTGAGCTT-----AGAGGGAA 1046
	/note="putative component of sperm tail outer dense fibers: tpx-1"	QY	876 ATGGTAGCCATGTTGATACAGTTGTTCAAATGTTCTCTGGATCTGCTT 935
	/codon_start=1	Db	1047 ACCTATAGGAGACTAGTGTAACTGTCGATCCAAATGACAGGCCTCTTCGGGA 1106
	/product="cysteine-rich secretory protein-2 CRISP-2"	QY	936 ATTTTACAAATATTTCTACAAATGGTAAAGAACAAATCTACACAC 995
	/protein_id="AAQ48090.1"	Db	1107 TTATATAGAATGTCCTTCATACAGCATTARGA -AAGTGTCTTGTAGGATACAC 1164
	/db_xref="CBI:5725526"	QY	996 TTGGATTTATATAAATCTGTGTTAATTCTGTGTTATAGGTGAA 1055
	/translation="MAFWQALFVGVLLPPTEQKDPDELTITNQIQREIATKHNLRQVSPGSNTIKMRWQVQANQKRNCLTHERSSTDRLINKTGENYMSDPTTSRRTVQSWYEEENFVQGPNSAVGHQYLVWSSFKVGVYQVNPDTIKYFVYCHQMGNTWAKSTPKYQOITPCASCPNNCDNGLCTNSCDPEDIILNSCSLK	Db	1165 TCTGGATTGTTGA---CCACTTGCTGCTCAATGTTGAGCAGTCAAGTGGAGAA 1221
BASE COUNT	458 a 301 c 299 g 373 t	QY	1056 TTGAAAGTTGATCTCATGACTGAGTCAAAACCTGGATGA 1107
ORIGIN	1 others	Db	1222 TTGGAAGTTGACCATAACTGGTCATCTGCTAGACTTTGAAAGGA 1273
Query Match	16.9%; Score 362.2; DB 10; Length 1432;	RESULT	15
Best Local Similarity	63.0%; Pred. No. 6.2e-57; Matches 675; Conservative 0; Mismatches 359; Indels 38; Gaps 6;	RATREG	
QY	37 CTGCAATGACATATTCCAGTGCTGTCCTGGTGTGGCTCTTCATCTTC 96	LOCUS	RATREG
Db	239 CAGCCATGGCTGGTCAGGATGCGTGTCTCTGTTCTACCA---TTC 295	DEFINITION	Rat epididymal glycoprotein (REG) mRNA, complete cds.
QY	97 CAGCAATGACATTAAGATCCGGTTACCTGCTTAACTCACCACACAGTC 156	VERSION	M3117.1 GI:202772
Db	296 CAGCCACAGAGGAAGGATCCAGACTTGCTGACTTGTACCAAACTTC 355	KEYWORDS	epididymal glycoprotein.
QY	157 AAGGGAGATTGTAATGCACTAGTGTACCTGGAGAGCAGTATCCTGCCAGAA 216	SOURCE	Rattus norvegicus (strain Sprague-Dawley) adult epididymus cDNA to mRNA.
Db	356 AAGAGAGATCACTGACACATGACTGGAGAGCACATGACCTGGAGCA 415	ORGANISM	Rattus norvegicus
QY	217 AATGTCGAAGTGGAAAGAGGCTGCAAGAACATGCCAAAGTGGCAACC 276	REFERENCE	Rattus. Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
	I.. 1432	1 (bases 1 to 1498)	

